WE CLAIM:

1. A system for caching network resources comprising:

a server having network resources stored thereon;

a client generating requests for the network resources;

an intermediary server configured to receive requests from the client and retrieve the network resources from the server;

a cache controlled by the intermediary server for caching selected network resources, wherein the cached resources include more than the requested resources and wherein at least some of the cached resources are selected both in response to the request and explicitly selected to prevent future client requests from being communicated to the server.

2. The system of claim 1 wherein the cache includes only a home page for at least one web site.

3. The system of claim 1 wherein the intermediary server comprises a front-end computer and a back-end computer.

4. The system of claim 3 wherein both the front-end computer and the back-end computer implement a cache data structure.

5. The system of claim 4 further comprising:

a first page cached on the front-end computer cache, the first page associated with a plurality of other resources, wherein the other resources are cached on the back-end computer cache.

6. The system of claim 5 wherein the association is explicit in links within the first page that point to the secondary resources.

7. The system of claim 5 wherein the association is implicit in user access patterns.

8. The system of claim 5 wherein the association is explicitly defined by the site owner.

9. The system of claim 1 wherein the cache is configured to store web pages and elements thereof.

10. The system of claim 1 wherein the cache is configured to store program constructs comprising software code, applets, scripts, active controls.

11. The system of claim 1 wherein the cache is configured to store files.

12. The system of claim 1 further comprising:

means within the intermediary server for merging a current request for network resources that are not in the cache with a prior issued pending request for the same network resources.

13. A cache system comprising:

a communication network;

a plurality of network-connected intermediary servers each having an interface for receiving client requests for network resources, each intermediary server having a cache associated therewith;

communication channels linking each intermediary server with a set of neighboring intermediary servers for exchanging cache contents amongst the intermediary servers.

14. A method for caching network data comprising:

communicating request-response traffic between two or more network-connected computing appliances;

implementing a cache coupled to the request-response traffic; and

selectively placing data from the request-response traffic into the cache at least partially based upon attributes of the client and/or server associated with the request-response traffic.

15. The method of claim 14 further comprising:

associating client attributes with the request-response traffic, the client attributes associating a relative priority with the traffic, wherein the act of selectively placing is at least partially based upon the client attribute.

16. The method of claim 14 further comprising:

associating client attributes with the request-response traffic, the client attributes associating a service level with the traffic, wherein the act of selectively placing is at least partially based upon the client attribute.

17. The method of claim 14 further comprising:

associating client attributes with the request-response traffic, the client attributes associating a service level with the traffic, wherein the act of selectively placing is at least partially based upon the a server assigned priority.

18. A cache system comprising:

a front-end server implementing a first cache and configured to receive client requests and generate responses to the client requests;

a back-end server implementing a second cache and configured to receive requests from the front-end server and generate responses to the front-end server;

an origin server having content stored thereon;

a communication channel linking the front-end server and the back-end server;

a cache management mechanism in communication with the front-end computer and the back-end computer to selectively fill the first and second caches.

19. The cache system of claim 18 wherein the cache management mechanism comprises a process within the front-end server for receiving responses to client requests and placing the received responses in the cache.

09835804.041501

20. The cache system of claim 18 wherein the cache management mechanism comprises a process within the front-end server for generating, sua sponte, requests and placing the responses to the sua sponte requests in the cache.

21. The cache system of claim 18 wherein the cache management mechanism comprises processes for populating one cache with contents from another cache.

22. A system for caching network resources comprising:

a plurality of intermediary servers configured to receive client requests and retrieve request-specified network resources;

a cache implemented within each of the intermediary servers and configured to store selected network resources;

a resolver mechanism for supplying a network address of the intermediary server to the client applications, wherein the resolver mechanism dynamically selects a particular intermediary server from amongst the plurality of intermediary servers based at least in part on the content of each intermediary server's cache.

23. The system of claim 22 further comprising:

a redirection mechanism within a first of the intermediary servers configured to redirect a client request from the first intermediary server to a second of the intermediary servers based at least in part on the content of the first and second intermediary server's caches.

09835580.041501

24. A cache system comprising:

a first front-end server implementing a first cache and configured to receive client requests and generate responses to the client requests;

a second front-end server implementing a second cache and configured to receive client requests and generate responses to the client requests;

an origin server having content stored thereon;

a communication channel linking the first front-end server and the second front-end server;

a cache management mechanism in communication with the first and second front-end computers to selectively fill the second cache in response to a client request received by the first front-end server.

25. The cache system of claim 24 wherein the cache management mechanism selectively updates the second cache based upon knowledge that subsequent client requests will be directed to the second front-end server.

25. The cache system of claim 24 wherein the cache management mechanism selectively updates the second cache based upon anticipation that subsequent client requests will be directed to the second front-end server.

27. A method of speculatively caching Internet content comprising:

receiving a current request for specified content;

obtaining the specified content in response to the current request; and

speculatively caching data in addition to the specified content.

28. The method of claim 27 wherein the act of speculatively caching data comprises determining data that is likely to be requested subsequent to the current request.

29. The method of claim 27 wherein the act of speculatively caching data comprises:

determining an ability for a server to respond to subsequent requests for the data; and

speculatively caching data when it is determined that the server's ability to respond to subsequent requests is less than a preselected level.